



# Data Quality Challenges in Systemic Lupus Erythematosus Trials: How Can This Be Optimized?

## Citation

Pike, Marilyn C., and Lexy Kelley. 2012. Data quality challenges in systemic lupus erythematosus trials: how can this be optimized? *Current Rheumatology Reports* 14(4): 324-333.

## Published Version

doi:10.1007/s11926-012-0261-7

## Permanent link

<http://nrs.harvard.edu/urn-3:HUL.InstRepos:10445642>

## Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA>

## Share Your Story

The Harvard community has made this article openly available.  
Please share how this access benefits you. [Submit a story](#).

[Accessibility](#)

# Data Quality Challenges in Systemic Lupus Erythematosus Trials: How Can This Be Optimized?

Marilyn C. Pike · Lexy Kelley

Published online: 13 May 2012  
© Springer Science+Business Media, LLC 2012

**Abstract** Major scientific advances in basic science, pharmacology, and translational medicine have allowed the discovery of new molecular targets whose manipulation by new chemical entities has led to treatments for inflammatory diseases, including rheumatoid arthritis, multiple sclerosis, and inflammatory bowel disease. Development of new agents for systemic lupus erythematosus (SLE) has lagged, however, because the protean manifestations of SLE present challenges for measuring therapeutic effects in a consistent manner. Composite end points combining several Disease Activity Indices (DAIs) are being used in ongoing global studies, but the uniform application of these complex DAIs across large numbers of clinical sites has proven difficult. We describe herein approaches that are being utilized to facilitate collection, review, and analysis of the clinical measures utilizing independent central adjudication committees.

**Keywords** Systemic lupus erythematosus · SLE · Data quality · Challenge · New agents · Therapy · Central adjudication committees · Disease Activity Index · Trials · Optimize

M. C. Pike  
Division of Rheumatology, Massachusetts General Hospital,  
Harvard Medical School,  
Boston, MA, USA

M. C. Pike (✉)  
MedPharm Consulting,  
8 Craigie Street,  
Cambridge, MA 02138, USA  
e-mail: mp@marilynpike.com

L. Kelley  
ReSearch Pharmaceuticals Services,  
520 Virginia Avenue,  
Fort Washington, PA 19034, USA  
e-mail: lkelley@rpsweb.com

## Introduction

Systemic lupus erythematosus (SLE) is an autoimmune disease with a wide array of clinical and laboratory manifestations that can affect every body system [1–3]. The pathogenesis of SLE involves genetic, immunologic, hormonal, and environmental factors, and current therapies include NSAIDs, antimalarials, corticosteroids, and immunosuppressive agents. Despite the use of non-target-specific, steroid-sparing, immunosuppressive agents such as methotrexate, azathioprine, mycophenolate mofetil, and cyclophosphamide, most patients require the ongoing use of low- to mid-dose levels of corticosteroids together with intermittent bursts of high doses [3–5]. The morbidity of these treatments prevents many of these patients from living productive lives, particularly those who are diagnosed in their early reproductive years. Patients with moderate to severe, remitting, relapsing disease would be better served by agents that target more specific immunologic molecular targets. The basic science, pharmacology, and translational medicine efforts for the discovery of therapies to alter new molecular targets have advanced significantly over the past 25 years; however, the protean nature of SLE has made it challenging to develop reproducible and scalable measures of disease activity that are applicable to global clinical trials requiring 150 to 300 investigative sites. Drug development for SLE has therefore lagged behind that of other inflammatory autoimmune diseases such as rheumatoid arthritis, multiple sclerosis, and inflammatory bowel disease [8–11].

Several Disease Activity Indices (DAIs) have been and are currently being used for the development of composite primary end points to test a number of therapeutics that are in clinical development for moderate to severe SLE. These DAIs include versions of the SLE Disease Activity Index (SLEDAI), including the original SLEDAI, which is

composed of 24 disease parameters heavily weighted toward the neurological and renal systems that are deemed “present or not present” within the previous 10 days of the assessment [12, 13]. The clinical parameters are given weighted scores of 1 to 8. The Safety of Estrogens in Lupus Erythematosus, National Assessment–Systemic Lupus Erythematosus Disease Activity Index (SELENA-SLEDAI), extended the original SLEDAI to include a flare index to measure moderate and severe flares over time in an SLE population whose disease was minimally active at baseline [14, 15]. The SLEDAI-2000, a modification of the original 24 SLEDAI clinical parameters, was developed to take into account persistent disease activity rather than only new or recurrent features of SLE [16]. It has been shown to be valid over both 30-day and 10-day periods [17].

Another commonly utilized DAI for SLE activity is the BILAG (British Isles Lupus Assessment Group) Index [18–20]. The BILAG is a group of physicians who have been meeting regularly since 1984, and originally produced what has become known as the “classic” BILAG Index, which consists of 86 clinical and laboratory parameters measured over a 28-day period. It is based upon a physician’s intention to treat and captures changes in active disease in different body systems simultaneously. The body systems included in the original version included constitutional, mucocutaneous, neuropsychiatric, musculoskeletal, cardiovascular, vasculitis, renal, and hematological. The classic BILAG underwent a major revision in 2004 to remove some damage items, increase the number of parameters to 97, add the gastrointestinal (GI) and ophthalmic systems, and include vasculitis terms within the appropriate body systems [21]. These parameters are scored by the treating physician as new (4), worse (3), same (2), improving (1), and not present (0). Within the nine body systems, grades of A through E are assigned based on a severity scoring index. While it is a complex instrument to use, it has been validated and shown to correlate with a physician’s intent to treat [22].

The SELENA-SLEDAI modification of the Physician’s Global Assessment, or MDGA, is a three-inch visual analogue scale (VAS) with superimposed landmarks ranging from 0 to 3 in which zero is equivalent to no disease activity, and 1, 2, and 3 correlate with mild, moderate, and the most severe disease possible, respectively. Note that an MDGA of 2.5 will trigger a severe flare by SELENA-SLEDAI criteria [14, 15]. Some studies have used a VAS of 100 mm, similar to that used in rheumatoid arthritis studies [23, 24•], which can also be successfully used in lupus trials with the SELENA-SLEDAI landmarks. The use of this instrument in clinical trials is dependent upon having sequential determinations performed by the same physician, as it is largely subjective.

In contrast to these activity indices, the Systemic Lupus International Collaborating Clinics/ACR (SLICC/ACR) is a

measure of damage or nonreversible change present for at least 6 months that is not related to ongoing active inflammation [25, 26].

The first drug to be approved for SLE in 60 years was belimumab, a monoclonal antibody that inhibits the activity of the B-lymphocyte stimulator (BLyS) [27, 28]. This drug was approved by the US Food and Drug Administration (FDA) in March 2011 [29]. The primary end point used in the two registration trials was the Systemic Lupus Response Index (SRI) [30•, 31•, 32•]. The SRI is a composite index composed of a  $\geq 4$ -point decrease in the SLEDAI, the absence of new flares in any of the eight body systems defined by no new A scores (representing severe disease activity), and no more than one new B score (representing moderate disease activity) in the classic BILAG, and no worsening ( $\geq 0.3$  increase) of the MDGA visual analogue three-point score [30•]. Using this composite end point, there was a modest but statistically significant increase in the response rate in patients receiving 10 mg/kg belimumab for 52 weeks [31•, 32•]. In the two trials that had identical entry criteria, BLISS-52 and BLISS-76 showed placebo response rates of 34 % or 44 %, and belimumab (10 mg/kg) response rates of 43 % and 58 % [31•, 32•]. This represented treatment effects of only 9 % and 14 %, respectively; however, all secondary end points demonstrated directional improvement in the treatment groups, the most important of which was a higher percentage of patients whose average prednisone dose was reduced by 25 % or more from baseline to  $\leq 7.5$  mg/d during weeks 40 through 52 [31•, 32•]. The main driver of the SRI was the 4-point or greater reduction in the SLEDAI score. Placebo rates of 34 % and 44 % may have been high because of the remitting and relapsing nature of the disease and because the patients were allowed increases in standard-of-care medications during the first part of the study, many of which can be quite effective.

Another targeted therapy currently in phase 3 is epratuzumab, which is a monoclonal antibody directed against the CD22 antigen on the B-cell surface [33–35]. Using the same DAIs, a different composite primary end point was independently developed and used in the EMBLEM phase 2 study [33]. The primary end point was the BILAG Based Composite Lupus Assessment (BICLA) at 12 weeks composed of BILAG 2004 improvement, defined as BILAG A’s at study entry improved to B/C/D and BILAG B’s at study entry improved to C/D and no BILAG worsening in other BILAG organ systems, no worsening in SLEDAI total score compared with study entry, no worsening in physician’s global VAS assessment of disease activity (defined as  $<10$  % increase) compared with study entry; moreover, patients who are treatment failures cannot be responders [33, 36•]. The results of the EMBLEM study showed that 21 % of patients receiving placebo met all criteria defined in the BICLA. Patients treated with epratuzumab in a total dose of

2,400 mg had a combined response rate of approximately 43 % at 12 weeks [33]. Although the treatment groups contained a relatively small number of patients (37–39 per arm), the lower placebo response resulted in the realization of a statistically significant, approximately twofold treatment effect.

The EMBLEM data were subjected to post-hoc analysis comparing the BICLA versus SRI composite end points [37••]. Application of the SRI to these data showed a placebo efficacy rate of greater than 50 %, or twice that observed when data were analyzed by BICLA criteria. No treatment effect was observed using SRI [37••]. Improvement using the BILAG-based BICLA requires a response in all body systems that were involved at baseline, as well as no new flares in the remaining body systems, which may account for the fact that a placebo response occurs less often. By contrast, if just one SLEDAI element worth at least four points resolves, while other features that were present at baseline stay the same or worsen slightly, then the patient could qualify as a responder in the SRI.

A similar analysis was applied to preliminary data obtained in the Biomarkers of Lupus Disease (BOLD) study, a study of 100 patients with SLE on immunosuppressive therapy [38••]. Half of the patients are undergoing withdrawal of their immunosuppressive therapy to determine whether this would be safe and lessen the time to flare. The placebo effect was examined using multiple outcome measures including the SRI-4, SRI-5, and a stringent end point similar to the BICLA, which did not, however, allow a responder to have even one B flare. These end points were compared with BOLD study protocol criteria minimally defined by either a  $\geq$ drop of one BILAG grade or a SLEDAI  $\geq$ 4-point reduction from baseline but anchored by the investigators' simple intent-to-treat-based determination, whether there was clinically significant improvement, no significant change, or clinically significant flare. A preliminary analysis at 4 and 8 weeks showed that the BICLA-like end point was superior to SRI in detecting improvement, and less likely to pick up a flare visit as improvement based on the BOLD standard [38••].

The potential for variability in the application of the SLEDAI, BILAG, and MDGA by physician evaluators worldwide poses a large challenge in multicenter clinical trials given the differences in disease expression in individual patients and the interpretation of the patient signs and symptoms. The focus of this article is on analyzing the potential pitfalls in the collection of clinical trial data in patients with SLE, describing approaches that have been used in the successful development of drugs for other clinical indications, and presenting solutions for assuring the uniformity of the data across multiple SLE clinical trial sites.

## Centralized Adjudication Committees for Complex Clinical End Points

Centralized adjudication committees (CACs) have been used successfully to analyze clinical end points that are not solely composed of objective laboratory data and are thus subject to variable and/or biased interpretation. These adjudication committees are different from data monitoring committees (also known as data safety monitoring committees or DMCs), which are a group of individuals established for large, randomized, multicenter trials in which a treatment is intended to reduce mortality or major cardiovascular events, recurrence of cancer, or other life-threatening events [39]. These committees are used in trials predominantly to monitor study treatments or procedures in which the results could possibly have a very favorable or highly unfavorable impact on safety, or in trials in which particularly fragile populations are involved and in whom the risk of mortality or other serious outcomes is high [39]. A DMC may be charged with determining whether trials should be aborted for futility if efficacy is found to be absent or if a detrimental effect on the population is discovered. Conversely, a trial may be stopped if the treatment is shown in an interim analysis to have a very favorable outcome such that it would be unethical to continue withholding therapy in a placebo-controlled trial.

In contrast to DMCs, CACs provide more detailed, specific data review, predominantly of efficacy end points [40, 41]. Given the investment of hundreds of millions of dollars required to bring a drug to late-stage development, and the intense level of audit and scrutiny applied to trial data before marketing approval of drugs and devices, most companies have employed centralized review and adjudication to assure bias-free interpretation of critical end points.

CACs are routinely used to analyze cardiac and pulmonary events for safety and efficacy [42, 43]. Criteria are developed prospectively for determining classification of events into myocardial infarction, acute coronary syndrome without infarction, appropriate application of angioplasty and stenting procedures, episodes of congestive heart failure, and respiratory deaths due to various causes [44, 45]. Studies have shown that there can be substantial discrepancies between evaluation of these events by the site investigator as compared with the central committee [43–45]. A recent analysis of clinical event classification (CHF and cardiogenic shock) for the Assessment of Pexelizumab in Acute Myocardial Infarction (APEX-AMI) trial showed that the clinical events committee agreed with site investigator assessments of CHF in only 45 % of CHF events, and 77 % of cardiogenic shock events [44].

In the Understanding the Potential Long-term Impacts on Function with Tiotropium (UPLIFT) trial, conducted in patients with chronic obstructive pulmonary disease

(COPD), a mortality adjudication committee determined cause-specific mortality due to respiratory, cancer, cardiovascular, sudden cardiac death, or sudden death, which was then compared with the site investigator assessment. In this study, there was complete agreement in only 50.2 %, incomplete agreement in 18.5 %, and no agreement in 31 % of 981 reported deaths. Thus, there is a high degree of variability in relying upon individual site assessments in classifying causes of what is usually considered a “hard” end point, mortality [45].

Assessments of complete and partial responses and time to progression in oncology trials are particularly difficult. These determinations are dependent upon physical examination and radiological imaging techniques, which are inherently variable. Central adjudication of these end points is critical for oncology drug development and provisional marketing approval in the absence of long-term mortality data [46, 47].

### Disease Activity Index Interpretation in Systemic Lupus Erythematosus Clinical Trials

The FDA guidance for the development of drugs for lupus specifies that sponsors should use an adjudication committee to determine whether patients meet prespecified criteria for responder status [48•]. The following describes two general approaches that can be utilized.

#### Post-Hoc Versus Ongoing Data Review and Adjudication in Systemic Lupus Erythematosus Trials

Adjudication can be performed when all patient visits are finalized and after site personnel have applied the DAIs to serial patient visits and assigned scores to BILAG, SLEDAI, and MDGA. The cumulative data are then provided to members of a CAC who have developed specific rules for verifying the correct application of the DAIs. All visits for a patient are reviewed and queries are issued to investigators to clarify the interpretation and correct application of the disease measurement tools. The advantage of this approach is that the adjudicator can assess the patient over the entire period of time being studied in sequential visits, which can help detect potentially illogical patterns in the data, but the disadvantage is that a significant amount of time has passed between the original patient visit and clinical adjudication/reassessment, leading to a less accurate recall of specific patient encounters by the investigator. In addition, CAC review of data following completion of all study visits prevents early identification of potentially incorrect application of the multiple DAIs, which can lead to an increase in protocol deviations and therefore a reduction in the *per protocol* population.

A preferable, albeit more labor-intensive approach is to conduct ongoing data review from the time of screening through database lock. This approach requires that patient data be made available to adjudicators soon after each patient visit, and therefore full-time coordination by data managers. Adjudicators also need to be available throughout the course of the trial rather than solely at the end. Given the history of failed or modestly successful development programs for lupus and the complexity of the disease and its assessments, we favor this approach, however. In particular, we encourage sponsors to include rigorous review of all patients screened for entry into the trial prior to randomization. Most of the trials currently recruiting, as well as those having been run in the recent past [24•, 31•, 48•], have strict entry criteria to ensure the following:

1. The patient has been correctly diagnosed as having lupus.
2. The patient has moderate to severe disease characterized by at least one A or two B BILAG 2004 body system scores, as well as a SLEDAI-2K score of 6 or greater. Most trials are now incorporating a minimum MDGA score to qualify for a study, as well as a prespecified type of immunosuppressive regimen.
3. The exclusion of subjects with severe neurological disease, active antiphospholipid antibody syndrome, and active lupus nephritis requiring cyclophosphamide, as these patients often require a different approach to therapy. While many investigators feel comfortable assessing the individual parameters in the SLEDAI-2K and BILAG 2004 in study patients, many do not have detailed experience in BILAG body system grading, especially the renal body system. Thus, it is important for external experts to verify the BILAG 2004 body system grades for study entry.
4. Another potential issue at screening and baseline visits that can greatly increase the placebo response rate using the SRI is the recording of inappropriate SLEDAI 2K eight-point items as “present.” All of these items, except for vasculitis, are neurologic features of SLE, most of which are exclusion criteria. In a post-hoc analysis of the EMBLEM study, Petri et al. [37•] compared the SRI with the BICLA using the same dataset and found a much higher placebo response rate using SRI compared with BICLA. In this analysis, 61 of 227 (27 %) differed in assignment of clinical response by SRI or BICLA, and the majority of these differences (47 SRI, but not BILAG responders) were explained by a disappearance of an eight-point item that was recorded at baseline: lupus headache ( $n=15$ ) or vasculitis ( $n=7$ ). The intended definition of lupus headache as applied in the original SLEDAI was that it was a manifestation of



**Table 1** Suggested correlation of SLEDAI-2000 and BILAG 2004 clinical parameters for SLE clinical trials<sup>a</sup>

SLEDAI-2000 term <sup>b</sup>	BILAG 2004 parameter <sup>c</sup>	Comment
Seizure	30. Seizure disorders 31. Status epilepticus	
Psychosis	24. Psychosis	
(3) Organic brain syndrome	23. Acute confusional state	
(4) Visual disturbance (limited to retinal changes of SLE)	69, 70. Posterior uveitis/retinal vasculitis—mild and severe 74. Retinal/choroidal vaso-occlusive disease 75. Isolated cotton-wool spots/cytoid bodies 76. Optic neuritis 77. Anterior ischemic optic neuropathy	SLEDAI-2000 definition includes only retinal lesions; original SLEDAI and SELENA- SLEDAI included all areas of the eye
(5) Cranial nerve disorder	27. Cranial neuropathy	
(6) Lupus headache	37. Lupus headache—severe unremitting 38. Headache from IC hypertension	Intended to be associated with active cerebral inflammation
(7) CVA	32. Cerebrovascular disease	
(8) Vasculitis	13. Major cutaneous vasculitis/thrombosis 18. Splinter hemorrhages 20. Cerebral vasculitis 51. Pulmonary hemorrhage/vasculitis 55. Coronary vasculitis 58. Lupus enteritis or colitis if vasculitis confirmed or other condition in which vasculitis is documented by imaging or biopsy	Confirm splinter hemorrhages are due to serious vasculitis
(9) Arthritis	41. Arthritis (severe) 42. Arthritis (moderate) if at least 2 joints involved	Note that SLEDAI-2000 requires ≥2 joints involved, while original SLEDAI and SELENA-SLEDAI required >2 joints (at least 3)
(10) Myositis	39. Myositis—severe 40. Myositis—mild	
(11) Urinary casts	88. Active urinary sediment and supporting laboratory value	Rarely detected in central laboratory urine specimens due to transport requirement; fresh urine sample in clinic required to detect
(12) Hematuria	88. Active urinary sediment and supporting laboratory value	
(13) Proteinuria >0.5 g/24 h, new or recent onset	81. Urine dipstick protein ≥3 83. UPCR ≥50 mg/mmol 85. Nephrotic syndrome	
(14) Pyuria >5 WBC/hpf, excluding infection	88. Active urinary sediment and supporting laboratory value	
(15) (New) rash	5. Skin eruption—severe 6. Skin eruption—mild 7. Angioedema—severe 8. Angioedema—mild 17. Periungual erythema/chilblains	
(16) Alopecia	15. Alopecia—severe 16. Alopecia—mild	
(17) Mucosal ulcers	9. Mucosal ulceration—severe 10. Mucosal ulceration—mild	
(18) Pleurisy	48. Pleurisy/pericarditis 50. Pleural effusion with dyspnea	
(19) Pericarditis	49. Cardiac tamponade 48. Pleurisy/pericarditis	
(20) Low complement	Not applicable to BILAG-2004	Not applicable to BILAG-2004
(21) Increased DNA binding	Central laboratory value	Not applicable to BILAG-2004
(22) Fever	1. Pyrexia, any 1–4	Note difference in glossary definitions
(23) Thrombocytopenia	Central laboratory value	
(24) Leukopenia	Central laboratory value	

<sup>a</sup> This is a guidance only. Depending on the clinical presentation, other BILAG-2004 terms may match with SLEDAI-2004 terms

<sup>b</sup> See Gladman et al. [16]

<sup>c</sup> See [20–22]

ongoing cerebral inflammation. Recent studies have shown that headache alone refractory to narcotics is no more common in the SLE population than in the population without lupus [49–52]. The wider availability of narcotics may falsely elevate the attribution of headache to SLE. Because the SLEDAI and BILAG are the key drivers of the SRI and the BICLA responses, it is critical to make certain that the clinical parameters contained in both that share similar, if not exactly the same, definitions are scored consistently across both DAIs at every patient visit. Table 1 shows examples of suggested data correlations for the DAIs used to measure SLE. Table 2 shows differences in the clinical and laboratory definitions between the BILAG 2004 and SLEDAI-2K glossaries. Note that the threshold for common laboratory abnormalities seen in SLE patients, such as leukopenia and thrombocytopenia, are different for SLEDAI-2K and BILAG 2004.

5. Another challenge in using the BILAG 2004 as a key end point is to ensure that site personnel assign the features of not present, improving, same, worse, and new in a logical fashion across different visits. There are very few clinical scenarios that would result in a change from “new,” “worse,” or “same” to “not present” in the 28-day period specified by the correct use of the BILAG 2004. Most parameters should go through an “improving” step before being scored “not present.” Why is this important? The BILAG 2004 Scoring Index is based on the intent-to-treat concept [21, 22]. Heavily weighted clinical parameters will produce an A or B score when “new,” “worse,” or “same” are marked, but reduction to “improving” reduces the body system

grade of an A parameter to B, and a B parameter to C. In many ongoing lupus phase 2 and phase 3 trials, swollen/tender joint counts and the Cutaneous LE Activity and Severity Index (CLASI) [53] are being used as exploratory outcome measures, but these tools can also provide additional data that can be utilized to verify that the SLEDAI-2K and BILAG are capturing joint and mucocutaneous parameters correctly. The joint count is the same 28 tender, swollen joint count used to capture active joints in rheumatoid arthritis assessments [54, 55], and the CLASI is an index that more precisely measures lupus skin involvement by applying weighted scores of activity (rashes, scale, hypertrophy) and damage (dyspigmentation) to segments of the body [53]. The CLASI is more heavily weighted toward areas of the body that are frequently involved in SLE, such as the face and scalp, while the Rule of Nines Burn Index [56], which is used in the BILAG 2004 to assess rash involvement, can greatly underestimate the degree of SLE skin involvement. Although both the joint count and CLASI give a perspective of these parameters only on the day of the visit, while SLEDAI and BILAG assess the previous 28 to 30 days, the CLASI and joint count may be particularly useful in identifying positive findings that are overlooked on the SLEDAI and BILAG. For example, if scalp erythema is scored as present and red on the CLASI, and the patient is noted to have alopecia, a central adjudicator may query a site to make certain that the term for “alopecia, severe” should be considered in addition to “alopecia, mild,” as erythema would indicate the possible characterization of inflammation associated with the alopecia.

**Table 2** Differences in SLEDAI-2000 and BILAG-2004 glossary definitions

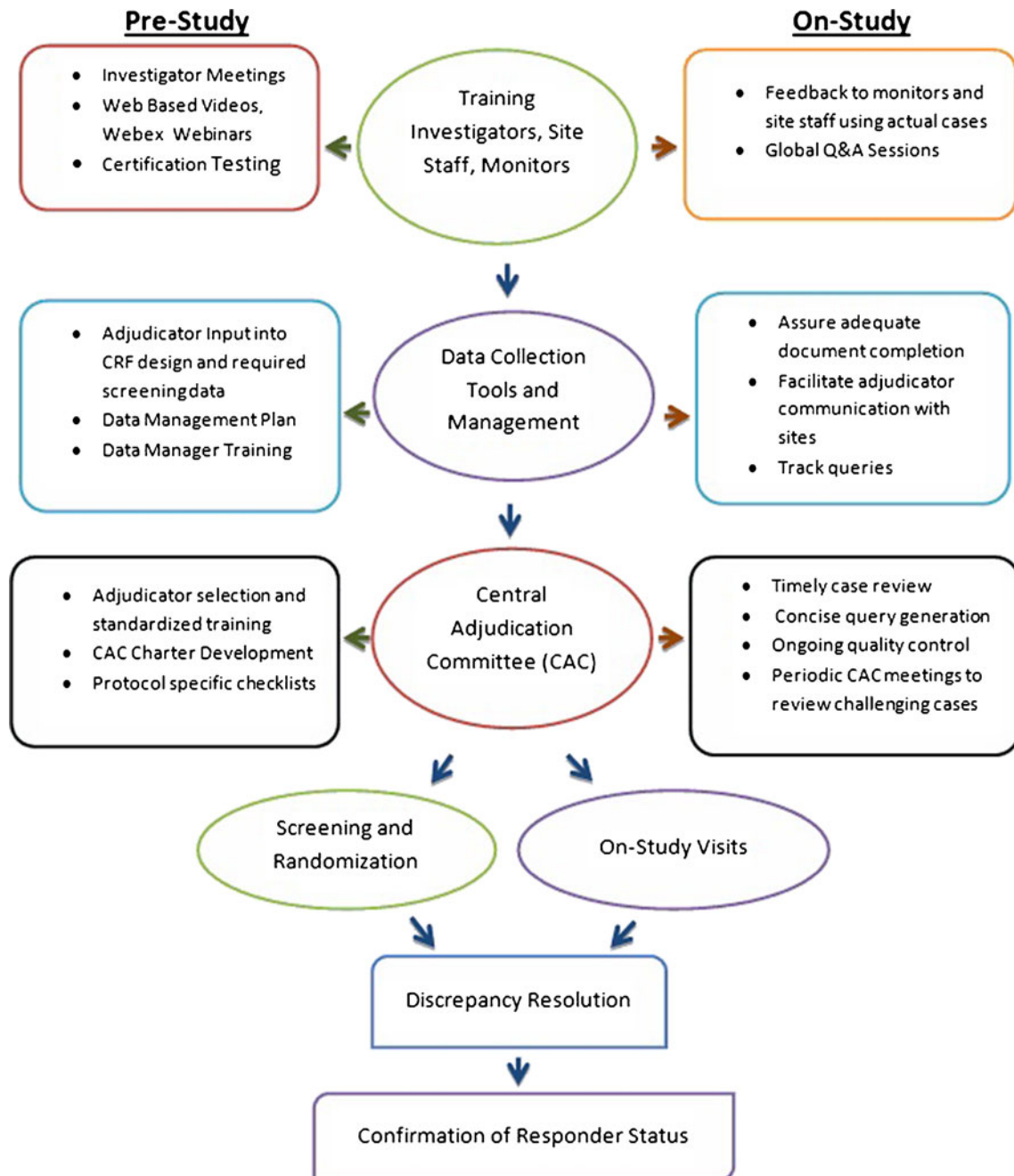
Parameter	BILAG-2004 [2–22]	SLEDAI-2K [16]
Ophthalmologic	Pathology in all areas of the eye	Pathology limited to retinal abnormalities
Arthritis	Minimum score requires 1 joint	Minimum score requires 2 joints
Fever	>37.5°C, 99.5°F	>38°C, 100.4°F
Proteinuria	Renal score B ≥50 mg/mmol=0.44 mg/mg=440 mg/g	>0.5 g/24 h=0.5 mg/mg=500 mg/g
Pleuritis		Requires objective evidence such as rub or supportive imaging
Pericarditis		Requires objective evidence such as rub or supportive imaging
Leukopenia	Hematology score C ≤3,900/mL considered abnormal	<3,000/mL abnormal
Thrombocytopenia	Score if ≤149,000	Score if <100,000
Anti-DNA and complement	Not included	Abnormalities included

## Operational Aspects of Central Adjudication for Global Systemic Lupus Erythematosus Trials

Although central review and adjudication of DAIs for consistency is desirable, and perhaps even essential to assure high-quality, interpretable data in this heterogeneous patient population, the execution of this process can be time consuming and costly for sponsors. The ideal components of this venture are included in Fig. 1.

## Investigator, Site Personnel, and Monitor Training and Certification

A uniform, concise, but detailed training program including SLEDAI, BILAG, correct use of MDGA, CLASI, and joint count consisting of lectures at Investigator Meetings, WebEx, and/or video training and certification testing (currently provided by the Lupus Foundation of America and others) that can be accessed through Web-based portals is



**Fig. 1** Operational aspects of central adjudication for global systemic lupus erythematosus. CRF, case report form



essential. The training and testing should be tailored to the specific study.

### Data Collection and Management

For large, global trials, Internet-based electronic data capture, case report forms, and source document templates should be designed with input from the adjudication team to facilitate collection of the correct data for the DAIs. The addition of preprogrammed edit checks minimizes manual queries by adjudicators and streamlines data review. An experienced data management team should be trained to streamline communication between adjudicators, sponsors, CRAs, and site personnel.

### Adjudicator Review of Screening Data Prior to Randomization

Screening source documents, subject narrative, and laboratory review should be done in a timely fashion to confirm that a subject has moderate to severe SLE by DAI entry criteria.

### On-Study Visits

On site source document verification by monitors should occur and all outstanding queries addressed before the adjudication team reviews the visits.

### Pre-database Lock Check for Discrepant Data

Trained data managers should review serial BILAG, SLE-DAI, and MDGA scores across all visits for each patient to identify unusual data patterns and have outliers reviewed by adjudicators.

### Confirmation of Response

Responder status should be confirmed by the CAC.

## Conclusions

The development of reproducible composite primary end points such as ACR20, 50, and 70 and DAS28-CRP and others for the measurement of changes in rheumatoid arthritis activity leads to more interpretable trial end points and a revolution in the development of drugs for treatment of this disease compared with 20 years ago [54, 56]. By studying the differential responsiveness of patients with different phenotypes and genotypes to the new rheumatoid arthritis therapies that target specific immunologic pathways, it will be possible to develop a better understanding of the pathogenesis and genetics of this disease.

SLE may represent several different diseases given the extremely variable clinical presentation and organ system involvement. The clinical development approach to SLE is more challenging than for rheumatoid arthritis; however, lessons have been learned from rheumatoid arthritis trials in constructing potentially viable clinical end points that can be used in large, global, multicenter trials involving more than 1,500 patients. While attempts are being made to develop and validate new versions of existing DAIs [57], the BILAG and SLEDAI are being used today in ongoing phase 2 and phase 3 trials. Given the potential subjectivity and complexity of applying these tools to patients with different SLE manifestations, it is essential to make certain that they are being applied uniformly and consistently. This paper has highlighted the importance of CACs for data review and confirmation of clinical responses in SLE clinical trials. In addition, we have described a process that is currently being utilized. While certain features such as grading of BILAG body systems can be automated with computer algorithms, as in the iBLIPs program developed by Isenberg et al. [19], and in certain other proprietary databases, which include automatic edit checks, it is nevertheless important to incorporate a thoughtful medical review to ensure consistency in the use of these DAIs in global clinical trials.

**Disclosure** MCP receives consultancy fees from UCB Pharma, Inc., MedImmune, Inc. and RPS, Inc. LK is a current employee of RPS, Inc. and a former employee of UCB Pharma, Inc.

## References

Papers of particular interest, published recently, have been highlighted as:

### •• Of major importance

1. Fessler BJ, Boumpas DT. Severe major organ involvement in systemic lupus erythematosus. Diagnosis and management. *Rheum Dis Clin North Am*. 1995;21:81.
2. Cervera R, Khamashta MA, Font J, Sebastiani GD, et al. Morbidity and mortality in systemic lupus erythematosus during a 10-year period: a comparison of early and late manifestations in a cohort of 1,000 patients. *Medicine*. 2003;82:299.
3. Panopalis P, Clarke AC. Systemic Lupus Erythematosus: clinical manifestations, treatment and economics. *Expert Rev Pharmacoeconomics Outcomes Res*. 2006;6:563–75.
4. Lo MS, Tsokos GC. Treatment of systemic lupus erythematosus: new advances in targeted therapies. *Ann N Y Acad Sci*. published online. 2012;1247:138–152.
5. Schur PH, Wallace DJ. Overview of the therapy and prognosis of systemic lupus erythematosus in adults. *UptoDate*. <http://www.uptodate.com> (accessed February 2012).
6. Aberer E. Epidemiologic, socioeconomic and psychosocial aspects in lupus erythematosus. *Lupus*. 2010;19:1118–24.

7. Baker K, Pope J. Employment and work disability in systemic lupus erythematosus: a systematic review. *Rheumatology*. 2009;48:281–4.
8. Smolen JS, Landewé R, Breedveld FC, et al. EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying anti-rheumatic drugs. *Ann Rheum Dis*. 2010;69:964.
9. Meng T, Weber MS, Hemmer B. Disease modifying agents for multiple sclerosis: recent advances and future prospects. *Drugs*. 2008;68:2445–68.
10. D'Haens G, Panaccione R, Higgins PDR, et al. The London Position Statement of the World Congress of Gastroenterology on biological therapy for IBD with the European Crohn's and Colitis Organization: when to start, when to stop, which drug to choose and how to predict response? *Am J Gastroenterol*. 2011;106:199–212.
11. Peyrin-Biroulet L, Deltenre P, de Suray N, et al. Efficacy and safety of tumor necrosis factor antagonists in Crohn's disease: meta-analysis of placebo-controlled trials. *Clin Gastroenterol Hepatol*. 2008;6:644.
12. Petri M, Hellmann D, Hochberg M. Validity and reliability of lupus activity measures in the routine clinic setting. *J Rheumatol*. 1992;19:53–9.
13. Bombardier C, Gladman DD, Urowicz MB, et al. Derivation of the SLEDAI: a disease activity index for lupus patients. *Arthritis Rheum*. 1992;35:630–40.
14. Petri M, Buyon J, Kim M. Classification and definition of major flares in SLE clinical trials. *Lupus*. 1999;8:685–91.
15. Buyon J, Petri M, Kim M. The effect of combined estrogen and progesterone hormone replacement therapy on disease activity in systemic lupus erythematosus: a randomized trial. *Ann Int Med*. 2005;142:953–62.
16. Gladman DD, Ibanez D, Urowicz MB. Systemic lupus erythematosus disease activity index 2000. *J Rheumatol*. 2002;29:288–91.
17. Tourma Z, Urowicz MB, Gladman DD. SLEDAI-2K for a 30-day window. *Lupus*. 2010;19:49–50.
18. Wallace DJ. Improving the prognosis of SLE without prescribing lupus drugs and the primary care paradox. *Lupus*. 2008;17:91.
19. Isenberg DA, Gordon C, BILAG Group. British Isles Lupus Assessment Group. From BILAG to BLIPS—disease activity assessment in lupus past, present and future. *Lupus*. 2000;9:651–4.
20. Hay EM, Bacon PA, Gordon C, et al. The BILAG index: a reliable and valid instrument for measuring clinical disease activity in systemic lupus erythematosus. *Q J Med*. 1993;86:447–58.
21. Isenberg DA, Rahman A, Allen E, et al. BILAG 2004. Development and initial validation of an updated version of the British Isles Lupus Assessment Group's disease activity index for patients with systemic lupus erythematosus. *Rheumatology*. 2004;2005(44):902–6.
22. Yee CS, Farewell V, Isenberg D, et al. The BILAG-2004 index is sensitive to change for assessment of SLE disease activity. *Rheumatology*. 2009;48:691–5.
23. Fitzgerald JD, Grossman JM, Kalunian KC. Annual systemic lupus erythematosus are rates in university practice [abstract]. *Arthritis Rheum*. 1998;41(9 Suppl):S219.
24. •• Wallace DJ, Kalunian KC, Petri MA, et al. Epratuzumab demonstrates clinically meaningful improvements in patients with moderate to severe systemic lupus erythematosus (SLE): results from EMBLEM™, a phase IIb study. *Arthritis Rheum*. 2010;62 (Suppl S10):S1452. *This was the first randomized trial using BICLA to show efficacy with epratuzumab in SLE.*
25. Gladman D, Ginzler E, Goldsmith C, et al. The development and initial validation of the Systemic Lupus International Collaboration Clinics/American College of Rheumatology Damage Index for systemic lupus erythematosus. *Arthritis Rheum*. 1996;39:363–9.
26. Lopez R, Davidson JE, Beeby MD, et al. Lupus disease activity and the risk of subsequent organ damage and mortality in a large lupus cohort. *Rheumatology*. Advance Access published November 21, 2011:1–8.
27. Cancro MP, D'Cruz DP, Khamashta MA. The role of B lymphocyte stimulator (BLyS) in systemic lupus erythematosus. *J Clin Invest*. 2009;119:1066–73.
28. Moore PA, Belvedere O, Orr A, et al. BLyS: member of the tumor necrosis factor family and B lymphocyte stimulator. *Science*. 1999;285:260–3.
29. FDA briefing information, belimumab (BENLYSTA), for the November 16, 2010 meeting of the Arthritis Advisory Committee. [www.fda.gov/AdvisoryCommittees/CommitteesMeetingMaterials/Drugs/ArthritisDrugsAdvisoryCommittee/ucm233578.htm](http://www.fda.gov/AdvisoryCommittees/CommitteesMeetingMaterials/Drugs/ArthritisDrugsAdvisoryCommittee/ucm233578.htm) (Accessed on November 16, 2010).
30. •• Furie RA, Petri MA, Wallace DJ, et al. Novel evidence-based systemic lupus erythematosus responder index. *Arthritis Rheum*. 2009;61:1143–51. *This was the first description of SRI for use as a primary composite end point in SLE clinical trials.*
31. •• Furie R, Petri M, Zamani O, et al. A phase III, randomized, placebo controlled study of belimumab, a monoclonal antibody that inhibits B lymphocyte stimulator, in patients with systemic lupus erythematosus. *Arthritis Rheum*. 2011;63:3918–30. *This was one of two phase 3 clinical trials submitted as key efficacy data leading to approval of belimumab in treatment of SLE.*
32. •• Navarra SV, Guzman RM, Gallacher AE. Efficacy and safety of belimumab in patients with active systemic lupus erythematosus: a randomised, placebo-controlled, phase 3 trial. *Lancet*. 2011;377:721–31. *This was one of two phase 3 clinical trials submitted as key efficacy data leading to approval of belimumab in treatment of SLE.*
33. Kalunian KC, Petri MA, Wallace DJ. BILAG-Measured improvement in moderately and severely affected body systems in patients with systemic lupus erythematosus (SLE) by epratuzumab: results from EMBLEM™. *Arthritis Rheum*. 2010;62(Suppl S10):S453.
34. Daridon C, Blassfeld D, Reiter K, et al. Epratuzumab targeting of CD22 affects adhesion molecule expression and migration of B-cells in systemic lupus erythematosus. *Arthritis Res Ther*. 2010;12: R204.
35. Traczewski P, Rudnicka L. Treatment of systemic lupus erythematosus with epratuzumab. *Brit J Clin Pharm*. 2011;71:175–82.
36. •• Wallace D, Strand V, Furie R, et al. Evaluation of treatment success in systemic lupus erythematosus clinical trials: development of the British Isles Lupus Assessment Group-based composite lupus assessment endpoint. *Arthritis Rheum*. 63(Suppl):S885. *This was the first description of the BICLA composite primary end point.*
37. •• Petri MA, Pike MC, Kelley L, et al. Systemic lupus erythematosus responder index assessment of responders in EMBLEM, a phase IIb study in patients with moderate to severe systemic lupus erythematosus. *Arthritis Rheum*. 2011;63(Suppl):S885. *This is a comparison of BICLA versus SRI as primary end points using the same clinical trial data.*
38. •• Sridharan S, Zhou T, Immermann F, et al. Low placebo responses and clinical components of the biomarkers of lupus disease (BOLD) study may provide useful insights for systemic lupus erythematosus clinical trial design. *Arthritis Rheum*. 2011;63(Suppl):S543. *This is a comparison of multiple end points using an SLE dataset.*
39. FDA guidance for clinical trial sponsors: Establishment and operation of clinical trial data monitoring committees. <http://www.fda.gov/downloads/RegulatoryInformation/Guidances/ucm127073.pdf> (Accessed February 8, 2012).
40. Kradjian S, Gutheil J, Baratelle AM, et al. Development of a charter for an endpoint assessment and adjudication committee. *Drug Inf J*. 2005;39:53–61.

41. Tyner CA, Somaratne R, Cabell CH, et al. Quintiles white paper executive summary: establishment and operation of clinical endpoint committees: best practice for implementation across the biopharmaceutical industry. <http://www.quintiles.com/information-library/white-papers/clinical-endpoint-committees> (Accessed February 8, 2012).
42. Mahaffey KW, Harrington RA, Akkerhuis M, et al. Systematic adjudication of myocardial infarction end-points in an international clinical trial. *Curr Control Trials Cardiovasc Med.* 2001;2:180–6.
43. Mahaffey KW, Harrington RA, Akkerhuis M, et al. Disagreements between central clinical events committee and site investigator assessments of myocardial infarction end-points in an international clinical trial: review of the PURSUIT study. *Curr Control Trials Cardiovasc Med.* 2001;2:187–94.
44. Mahaffey KW, Wampole JL, Stebbins A, et al. Strategic lessons from the clinical event classification process for the assessment of pexelizumab in acute myocardial infarction (APEX-AMI) trial. *Contemp Clin Trials.* 2011;32:178–87.
45. McGarvey LP, Magder S, Burkhart D, et al. Cause-specific mortality adjudication in the UPLIFT COPD trial: findings and recommendations. *Resp. Med. Advance Access.* 2011, doi:10.1016/j.rmed.2011.10.009.
46. Amit O, Bushnell W, Dodd L, et al. Blinded independent central review of the progression-free survival endpoint. *Oncologist.* 2010;15:492–5.
47. FDA Guidance for industry: clinical trial endpoints for the approval of cancer drugs and biologics. May 2007, <http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm071590.pdf> (Accessed February 8, 2012).
48. •• FDA Guidance for Industry: Systemic lupus erythematosus-Developing medical products for treatment. June 2010, [http://www.elsevierbi.com/~media/Images/Publications/Archive/The%20Pink%20Sheet%20Daily/2010/9/14/14100914002/0621\\_sle\\_guidance.pdf](http://www.elsevierbi.com/~media/Images/Publications/Archive/The%20Pink%20Sheet%20Daily/2010/9/14/14100914002/0621_sle_guidance.pdf) (Accessed February 8, 2010). *This is the regulatory guidance for designing SLE trials.*
49. Schur PH, Khoshbin S. Neurologic manifestations of systemic lupus erythematosus. Disclosures: Last literature review version 19.2: May 2011. (This topic last updated: February 18, 2011).
50. Jennekens FG, Kater L. The central nervous system in systemic lupus erythematosus. Part 1. Clinical syndromes: a literature investigation. *Rheumatology.* 2002;41:605.
51. Omdal R, Henriksen OA, Mellgren SI, Husby G. Peripheral neuropathy in systemic lupus erythematosus. *Neurology.* 1991;41:808.
52. Mitsikostas DD, Sfikakis PP, Goadsby PJ. A meta-analysis for headache in systemic lupus erythematosus: the evidence and the myth. *Brain.* 2004;127:1200.
53. Bonilla-Martinez ZL, Albrecht J, Troxel AB, et al. The cutaneous lupus erythematosus disease area and severity index. *Arch Dermatol.* 2008;144:173–80.
54. Aletaha D, Nell VP, Stamm T, et al. Acute phase reactants add little to composite disease activity indices for rheumatoid arthritis: validation of a clinical activity score. *Arthritis Res Ther.* 2005;7:R796–806.
55. Wachtel TL, Berry CC, Wachtel EE, Frank HA. The inter-rater reliability of estimating the size of burns from various burn area chart drawings. *Burns.* 2000;26:156–70.
56. Lillegraven S, Prince FHM, Shadick NA, et al. Remission and radiographic outcome in rheumatoid arthritis: application of the 2011 ACR/EULAR remission criteria in an observational cohort.
57. Touma Z, Urowitz MB, Fortin PR, et al. Systemic lupus erythematosus disease activity index 2000 responder index-50: a reliable index for measuring improvement in disease activity. *J Rheumatol.* 2011;38:868–73.